

# **ARTICLE**

Received 21 Nov 2013 | Accepted 28 May 2014 | Published 15 Jul 2014

DOI: 10.1038/ncomms5231

# Substrate-bound outward-open state of the betaine transporter BetP provides insights into Na<sup>+</sup> coupling

Camilo Perez<sup>1,\*,†</sup>, Belinda Faust<sup>1,\*,†</sup>, Ahmad Reza Mehdipour<sup>2,\*</sup>, Kevin A. Francesconi<sup>3</sup>, Lucy R. Forrest<sup>2,†</sup> & Christine Ziegler<sup>1,4</sup>

The Na<sup>+</sup>-coupled betaine symporter BetP shares a highly conserved fold with other sequence unrelated secondary transporters, for example, with neurotransmitter symporters. Recently, we obtained atomic structures of BetP in distinct conformational states, which elucidated parts of its alternating-access mechanism. Here, we report a structure of BetP in a new outward-open state in complex with an anomalous scattering substrate, adding a fundamental piece to an unprecedented set of structural snapshots for a secondary transporter. In combination with molecular dynamics simulations these structural data highlight important features of the sequential formation of the substrate and sodium-binding sites, in which coordinating water molecules play a crucial role. We observe a strictly interdependent binding of betaine and sodium ions during the coupling process. All three sites undergo progressive reshaping and dehydration during the alternating-access cycle, with the most optimal coordination of all substrates found in the closed state.

<sup>&</sup>lt;sup>1</sup> Structural Biology Department, Max-Planck Institute of Biophysics, Frankfurt am Main 60438, Germany. <sup>2</sup> Computational Structural Biology Group, Max-Planck Institute of Biophysics, Frankfurt am Main 60438, Germany. <sup>3</sup> Institute of Chemistry—Analytical Chemistry, University of Graz, Graz 8010, Austria. <sup>4</sup> Institute of Biophysics and Biophysical Chemistry, University of Regensburg, Regensburg 95053, Germany. \* These authors contributed equally to this work. † Present addresses: Institute of Molecular Biology and Biophysics, Eidgenössische Technische Hochschule (ETH) Zürich, CH-8093 Zürich, Switzerland (C.P.); Structural Genomics Consortium, University of Oxford, Oxford OX3 7DQ, UK (B.F.); Computational Structural Biology Unit, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD 20852, USA (L.R.F.). Correspondence and requests for materials should be addressed to C.P. (email: camilo.perez@mol.biol.ethz.ch) or to L.R.F. (email: lucy.forrest@nih.gov) or to C.Z. (email: christine.ziegler@biophys.mpg.de).

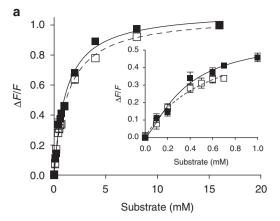
etP, an osmotic-stress-regulated transporter from the betaine/choline/carnitine transporter family shares the conserved LeuT-like fold1 of two inverted structural repeats of five transmembrane helices<sup>2,3</sup> (Supplementary Movie), whose pseudo-symmetry has been proposed as the mechanistic basis of the alternating-access mechanism of transport<sup>4</sup>. Various transporters from diverse sequence families share this highly conserved fold<sup>5</sup>, including other sodiumcoupled transporters such as the hydantoin transporter Mhp1<sup>6</sup> and a galactose transporter vSGLT<sup>7</sup>. Recently, several crystal structures of conformationally asymmetric BetP trimers have been reported<sup>2</sup>, which almost completed our view of the alternating-access cycle for BetP. Nevertheless, some fundamental questions were left unanswered. In particular, a substrate-bound outward-open state of BetP was missing and therefore the molecular understanding of the key events of substrate and co-substrate binding, and the subsequent transition to the closed state, was unclear. Recent structural work on BetP identified a functional important mutant (G153D) carrying an aspartate in the flexible stretch of transmembrane helix (TMH) 3, the first helix in the first repeat in LeuT-like fold transporters, which we denominate in the following as TMH1'. G153D is able to transport choline<sup>8</sup> (Supplementary Movie), exhibit an elevated apparent affinity for Na+ and preferably crystallizes in an outward-facing state<sup>2</sup>. We have reported that crystallization of this mutant resulted in outward-facing open and outwardoccluded apo states<sup>2</sup>.

As most of the BetP structures exhibit only moderate resolution of  $\sim 3\,\text{Å}$ , at which an unambiguous assignment of substrates to the electron density is hampered, here we cocrystallize BetP-G153D in the presence of a choline analogue containing arsenic instead of nitrogen, with which the substrate can be traced by arsenic anomalous scattering. These types of organo-arsenic compounds are natural constituents of many organisms, particularly from marine environments<sup>9</sup>. We show here that the slightly altered chemical nature of arseno-choline does not affect binding to BetP-G153D, although it leads to a reduction in Km compared with choline.

Co-crystallization of BetP-G153D and arseno-choline result in crystals in which BetP adopts the missing substrate-bound outward-facing state, with arseno-choline bound to the central binding site and revealing the binding mode of choline in detail. Subsequently, we use the new outward-facing substrate-bound structure in molecular dynamics (MD) simulations to compare the detailed binding of substrate and sodium in outward-facing open ( $C_e$ ), and outward-facing substrate-bound ( $C_e$ S) states to that in the closed state ( $C_c$ S)<sup>2,10</sup>. The results indicate that the two sodium ion-binding sites undergo major re-arrangements that affect their degree of hydration. Thus, the new  $C_e$ S state represents a missing puzzle piece in what is now an unprecedented set of structural snapshots for a secondary transporter, that together emphasize the role of both sodium ions in shaping and finally closing the central substrate-binding site in accordance with the induced fit model.

## **RESULTS**

Binding and transport properties of arseno-choline. Titration of BetP-G153D reconstituted in proteoliposomes with arseno-choline in the presence of 100 mM NaCl (Fig. 1a) revealed a similar binding affinity to that of choline (Supplementary Table 1). The arseno-choline-binding isotherm reveals biphasic-binding behaviour and positive cooperativity ( $n_{\rm Hill}=1.69\pm0.08$ ), similar to the binding behaviour of choline. Arsenic sustains two additional electron shells relative to nitrogen. Thus, the coordination of the trimethylarsonium group may differ slightly from



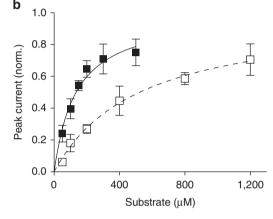


Figure 1 | Binding and transport properties of arseno-choline. Closed squares (choline), open squares (arseno-choline). (a) Tryptophan fluorescence-binding curve of choline and arseno-choline to BetP-G153D reconstituted in proteoliposomes. Each point shows the average for eight individual measurements. The error bars represent s.d. (b) Normalized peak currents of sodium-coupled transport of choline and arseno-choline to BetP-G153D. The graphs show average values from three individual recordings and the corresponding s.d.

that of trimethylammonium, which involves cation- $\pi$  interactions with tryptophan residues in the substrate-binding site<sup>2</sup>. However, the substrate hydroxyl group appears to interact with the carboxylic side chain of Asp153, which may suffice to compensate for any differences owing to the As atom so that the binding affinity is comparable to that of choline. To demonstrate transport of arseno-choline, solid-supported membrane (SSM)-based electrophysiology was performed with proteoliposomes of reconstituted<sup>10</sup> BetP-G153D. A transient current corresponding to transmembrane charge displacement was detected after activation with arseno-choline, showing that this compound can be transported (Fig. 1b). The peak maximum current plotted as a function of substrate concentration showed a fourfold increase in  $K_{\rm m}$  for arseno-choline ( $K_{\rm m} = 523 \pm 70 \,\mu{\rm M}$ ) compared with that of choline ( $K_{\rm m} = 135 \pm 28 \,\mu{\rm M}$ ) (Fig. 1b; Supplementary Table 1), suggesting some effect of the larger trimethylarsonium group on the transport kinetics of this substrate. Nevertheless, our thermodynamic and kinetic studies of arseno-choline indicate that the arseno derivative can be considered a suitable substrate analogue. Therefore, we included this compound in co-crystallization trials, allowing us to trace binding to BetP-G153D using arsenic anomalous scattering.

An arseno-choline-bound outward-open state (C<sub>e</sub>S). Co-crystallization of BetP-G153D with arseno-choline yielded crystals

diffracting to 2.95 Å resolution (Supplementary Table 2). As observed previously for this mutant, the crystal structure reveals an asymmetric trimer<sup>2</sup>, although unlike the earlier crystal structure, chain A and B adopt an arseno-choline-bound outward-open conformation (C<sub>e</sub>S), and chain C an arsenocholine-bound inward-open conformation (CiS) (Fig. 2a). The previous crystal structure of BetP-G153D contained substrate-free outward-occluded ( $C_{eoc}$ ) and outward-open states ( $C_e$ )<sup>2</sup>. The appearance of the outward-facing state found exclusively in crystals of this mutant was attributed to the sixfold higher apparent affinity for sodium, which presumably stabilizes this conformation by occupation of the Na2 site<sup>11,12</sup>. The new CeS state that we report here is also not frequently observed in the LeuT-like fold family. Indeed, only structures of the amino-acid antiporter AdiC have been observed in a similar C<sub>e</sub>S state<sup>13</sup>, while for LeuT only the competitive inhibitor-bound outward-open state<sup>14</sup> has been reported (Supplementary Fig. 1).

In this study, we used an oversaturating concentration of arseno-choline (approximately eightfold above  $K_d \sim 1 \, \text{mM}$ ), which appears to preferentially populate the C<sub>e</sub>S state, presumably owing to the distinct binding and transport characteristics of this substrate. Thus, the structures of protomers A and B, which we assign to the C<sub>e</sub>S state, contain a fully open central cavity that renders the central substrate and co-substrate-binding sites sterically accessible from the periplasm, whereas a  $\sim 18\text{-Å-thick}$  protein density occludes these sites from the cytoplasm (Fig. 2a). By contrast, in protomer C the central binding site is fully accessible to the cytoplasm, and is therefore assigned to the C<sub>i</sub>S state (Fig. 2a).

An anomalous difference Fourier map revealed prominent peaks in the central binding sites of chains A, B and C (Fig. 2b). These could be observed when using a contour level of up to  $6.1\sigma$  for chain A,  $11.8\sigma$  for chain B and  $5.2\sigma$  for chain C. No additional arsenic anomalous peaks were observed at comparable contour levels besides the ones in the central binding sites, indicating that, at least in these conformations of G153D, other high affinity arseno-choline-binding sites could not be discerned.

Similar to what was observed for the C<sub>c</sub>S state<sup>1</sup>, the binding site in the C<sub>c</sub>S state is formed by three tryptophan residues from

TMH6' arranged as in a prism (Trp-prism); Trp373, Trp374 and Trp377 coordinate the trimethylarsonium group of the substrate, while the carboxyl group of Asp153 and the backbone amide of Gly151 appear to hydrogen bond with the hydroxyl group of arseno-choline (Fig. 2b). During multiple 200-ns-timescale MD simulations of CeS monomers of BetP-G153D embedded in a hydrated bilayer of POPG lipids, these hydrogen bonds are readily formed, and the same overall coordination is maintained throughout (Fig. 3), providing confidence in the choice of arsenocholine as a substrate analogue. A similar interaction between a side-chain carboxyl group and the hydroxyl group of choline has been observed in other choline-binding proteins, including: the substrate-binding protein ChoX of the ABC transporter ChoXWU from *Escherichia coli*<sup>15</sup>; the major autolysin (C-Lyta) from Streptococcus pneumoniae<sup>16</sup>; and the choline-binding protein OpuBC from Bacillus subtilis<sup>17</sup>. This similarity indicates parallels in the binding modes of choline in membrane and soluble proteins.

This position of the arseno-choline in the  $C_eS$  conformations differs from that in the  $C_iS$  state, where the trimethylarsonium group is displaced from the Trp-prism (Fig. 2b, right). The latter is similar to what was observed for choline as well as for betaine in the  $C_iS$  state of a previously reported structure<sup>8</sup>.

Analysis of the B-factors (Supplementary Fig. 2) and root mean-squared fluctuations (Supplementary Fig. 3) of the substrate-binding site residues in the CeS state suggest that the unwound segment of TMH1' and perhaps also residues Trp373 and Trp377 of TMH6' that form the Trp-prism, increase their conformational rigidity upon substrate binding, as compared with the Ce state of BetP (Supplementary Fig. 2). Although such a comparison of B-factors should be considered with caution because the structures differ, a similar trend is observed on comparison of the related states for AdiC and LeuT (Supplementary Fig. 2) for the unwound segments of TMH1' and for TMH6', in which B-factors are reduced for the residues involved in substrate or competitive inhibitor binding, respectively<sup>13,14</sup>. Such entropic changes could conceivably facilitate the conformational changes between outward- and inward-facing states, imposing a dependence of the transition on substrate and co-substrate binding.

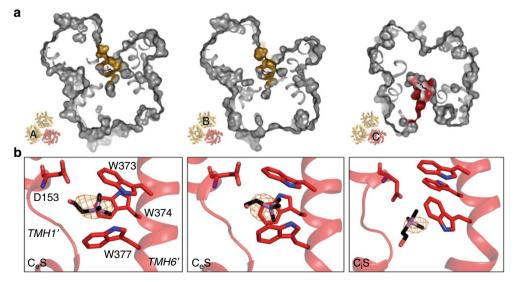
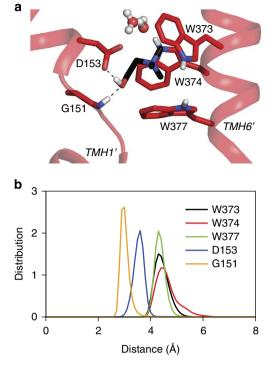


Figure 2 | Conformational states and arseno-choline binding observed in BetP-G153D. (a) Surface representation showing the periplasmic cavity. Chains A and B are in substrate-bound outward-open state ( $C_e$ S), while chain C is in substrate-bound inward-open conformation ( $C_i$ S). (b) Central binding sites of chains A, B and C, respectively. Arseno-choline is shown in black, red and purple sticks. Anomalous difference Fourier maps are shown at  $4.0\sigma$ ,  $7.0\sigma$  and  $4.0\sigma$  levels for substrates in chain A, B and C, respectively.



**Figure 3 | C<sub>e</sub>S state coordination of choline in the S1-binding site of BetP-G153D. (a)** Typical coordination of choline (black sticks) during the MD simulations by TMH1' and TMH6' (cartoon helices). Key residues (sticks) and water molecules (ball-and-stick) are shown. Non-covalent interactions are shown as dashed lines. (b) Distances were measured between the N atom of choline and the center of mass of the side chains in the tryptophan box (W377, W374 and W377), or between the hydroxyl oxygen of choline and the N backbone atom of G151 or the closest carboxyl oxygen atom of D153. The distribution was calculated over three trajectories, each 200 ns long.

The Na2 sodium-binding site in the CeS state. NaCl titration of BetP-G153D reconstituted in proteoliposomes revealed positive cooperativity  $(n_{Hill} = 1.9 \pm 0.2)^2$ , similar to that observed for wildtype (WT) BetP<sup>2</sup>. This cooperativity suggests that the mutant retains the two-Na<sup>+</sup> binding stoichiometry of WT BetP<sup>10</sup>. Recently, two sodium-binding sites in BetP were localized and functionally validated based on a crystal structure of the CcS state, symmetry considerations and molecular simulations<sup>10</sup>. The Na2 site is highly conserved among symporters of the LeuT-like superfamily<sup>18</sup>, and a positive peak in the Fo-Fc electron-density map was observed at this site in the earlier BetP C<sub>c</sub>S structure<sup>2</sup>. We detected an equivalent peak in the Fo-Fc electron-density map at the Na2 site in the new CeS state structures (Fig. 4a), suggesting that sodium is bound to the Na2 site in the substratebound outward-open (CeS) conformation. This proposal is supported by the fact that a Na + was stably coordinated at this Na2 site during multiple 200-ns-long MD simulations of the BetP-G153D C<sub>e</sub>S state (Fig. 4b-e).

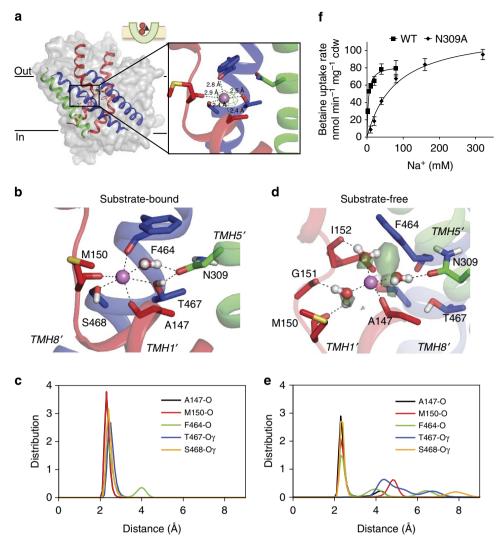
We note that the Na2 site coordination in this outward-open state differs slightly from that in structures of the closed C<sub>c</sub>S state<sup>2</sup>. For example, there are occasional fluctuations in the coordination by the F464 backbone carbonyl of C<sub>e</sub>S (Fig. 4e). Interestingly, the coordination of the Na2 site involves Asn309 from the gating helix TMH5′ (Fig. 4c,d). Asn309 is at the equivalent position of the key cation-substituting group Arg262 in a sodium-independent antiporter, CaiT<sup>19</sup>, and is one turn above the position equivalent to Asp189 in vSGLT, which is thought to interact with the Na2 ion in that transporter<sup>20,21</sup>.

We tested the importance of this residue for BetP by replacing Asn309 with alanine. Uptake assays in *E. coli* MKH13 cells reveal that this substitution increases the apparent  $K_{\rm m}$  of BetP for sodium  $\sim$  20-fold (Fig. 4f; Table 1) compared with that of BetP WT, confirming the predicted role of Asn309 in the sodium dependence of substrate transport, and suggesting that the interaction observed in the  $C_{\rm e}$ S simulations is physiologically important.

Unlike the aforementioned CeS- and CcS-state structures of BetP, no positive peak was detected at the Na2 site in the substrate-free outward-facing state<sup>2</sup> (C<sub>e</sub>). This lack of a positive peak is consistent with the observed distortion of the binding site in the C<sub>e</sub> structure (Table 2) owing to a number of subtle changes in the structure. These changes are: a ~2-Å displacement and unwinding of the middle of TMH8' near Phe464; a ~2-Å vertical translation of the cytoplasmic half of TMH3' towards the site; as well as a shift in the unwound segment of TMH3' away from the site. As a result, Thr467 and Met150 appear no longer close enough to coordinate the ion in the C<sub>e</sub> state compared with the C<sub>e</sub>S state (Fig. 5a,b). Indeed, during simulations of C<sub>e</sub> structures, an ion in the Na2 region was unable to interact with all ligands simultaneously, and mainly interacted with Ala147, Phe464 and Ser468 (Fig. 4d,e). Nevertheless, the ion did not escape from the Ce-Na2 site during these simulations, unlike ions placed in the Na2 site of inward-facing (C<sub>i</sub>) states of other transporters such as vSGLT<sup>22</sup> and Mhp1<sup>6</sup>, where all contributing oxygen atoms are > 2 Å further apart than in the  $C_e$  state (Table 2), so that the site is effectively abolished. Instead, in simulations of C<sub>e</sub>, the Na <sup>+</sup> ion maintained coordination by one or other backbone oxygen of TMH3' (Fig. 4d), as well as an average of  $2.0 \pm 0.8$  water molecules within 3 Å of the ion (Fig. 3e). The high degree of hydration of the Ce-Na2 site is consistent with the lack of a clear peak. This behaviour is distinct from ions at Na2 in the inward C<sub>i</sub>S or closed C<sub>c</sub>S states, where the site is fully disrupted (and solvent accessible) or fully formed (and completely inaccessible), respectively<sup>10</sup> (Table 2).

In summary, these results suggest that the coordination of the Na2 site becomes progressively optimized as the protein transitions from the  $C_{\rm e}$  through the  $C_{\rm e}S$  to the  $C_{\rm c}S$  states, where it adopts the most optimal coordination.

The Na1' sodium-binding site in the CeS and Ce states. The position of the second sodium-binding site was recently proposed based on the structural symmetry between repeats 10, and the proposal for this so-called Na1' site (which is distinct from the Na1 site in LeuT) was validated using biophysical assays and molecular simulations<sup>10</sup>. That study indicated a sodium ion at this position in the CcS state would be coordinated by the hydroxyl groups of Thr246 and Thr250 from TMH3', by a water molecule that hydrogen bonds to Thr499 from TMH9', and potentially by a cation- $\pi$  interaction to the aromatic ring of Phe380 from TMH6'. Although no positive peak for a Na1' sodium ion was observed in the *Fo-Fc* electron-density map of the outward-facing CeS state (Fig. 6a), molecular simulations of this C<sub>e</sub>S state indicated that an ion at the Na1' site interacts with the two Thr residues in TMH3' (Fig. 6b) in the same manner as in the closed C<sub>c</sub>S state<sup>10</sup>, albeit with the ion further from the Phe380 ring  $(5.1 \pm 0.5 \,\text{Å}$  on average in the C<sub>e</sub>S state) than in the closed  $C_cS$  state<sup>10</sup> (4.6 ± 0.3 Å). Water molecules became inserted between Phe380 and the ion in the simulations of the CeS state (Fig. 6b) because the extracellular halves of TMH3' and TMH6' are further apart (Fig. 5d) than in the closed C<sub>c</sub>S state (Table 2). These water molecules exchanged readily with bulk waters through the extracellular pathway (Fig. 7). Specifically,  $\sim 70$ different water molecules came within 3 Å of the Na1' ion in the



**Figure 4 | The Na2-binding site in outward-open states of BetP-G153D.** (a) The  $C_eS$  state structure (protomer B) in side view with Na2 and adjacent central substrate-binding site in stick representation, coordinating Na<sup>+</sup> (violet sphere) and arseno-choline (black, red and purple sticks), respectively. The inset shows the coordination of a Na<sup>+</sup> ion (violet sphere) in the Na2 site. The  $F_o - F_c$  electron density map is shown in green at a level of  $3.0\sigma$ . (**b-e**) The Na2 site in simulations of the substrate-bound  $C_eS$  (**b,c**) and substrate-free  $C_e$  (**d,e**) states: (**b,d**) Example configurations in simulations of the  $C_eS$  state (**b**) and the  $C_eS$  state (**d**). Proteins and ions are presented as in (**a**), waters are shown as balls-and-sticks and water densities are shown as green surfaces. (**c,e**) Distance between Na<sup>+</sup> ion and Na2-site oxygen atoms during three molecular dynamics simulations of the substrate-bound  $C_eS$  state (**c**), each 200 ns long, and the substrate-free  $C_e$  state (**e**), each 100 ns long. (**f**) Betaine uptake rates in nmol min<sup>-1</sup> mg<sup>-1</sup> cell dry weight (cdw) were measured as a function of the external sodium concentration in *E. coli* MKH13 cells expressing BetP WT or the mutant N309A. Each point shows the average of at least three independent experiments. Error bars represent s.d.

Table 1 | Kinetic parameters for apparent sodium affinity measured from betaine uptake of WT BetP and N309A mutant.

	<i>Κ</i> <sub>m</sub> μΜ	V <sub>max</sub> (nmol min <sup>– 1</sup> mg <sup>– 1</sup> cdw)		
WT	3.8 ± 0.9*	82.9 ± 4.8		
N309A	79.7 ± 16.4	123.0 ± 9.7		

cdw, cell dry weight; WT, wild type.

\*Each value corresponds to the average of at least three independent experiments. ± corresponds to s.d.

outward-facing  $C_eS$  state and remained for at least 50 ps. Nevertheless, the ion remained coordinated to Thr246 and Thr250 in these simulations (Fig. 6c), apparently stabilized by a

network of hydrogen bonds formed by, on average,  $2.2 \pm 0.4$  water molecules residing within 3 Å of the Na1' ion.

We also compared the Na1′ site in the  $C_eS$  state (Fig. 6a–c) with that in the  $C_e$  state (Fig. 6d,e), and found a similar situation to that described above for the Na2 site. That is, the sodium ion lacked a typical coordination shell in the  $C_e$  state, and yet remained coordinated to one or more of the same ligands (the side-chain hydroxyl of Thr250 in this case), with all other ligands then substituted by water molecules (Fig. 6d) that were exchanging with the extracellular solution (Fig. 7). These arrangements were formed within 40 ns in three different simulations of the  $C_e$  state, after initially placing the ion centrally between the binding residues. The altered arrangement of the  $C_e$  state appears to be attributable to a  $\sim$  3-Å vertical translocation of TMH9′ relative to TMH3′ (Fig. 5c), which separates Thr499 from the other residues in the site (Figs 5c and 6d), preventing formation of a stable water bridge that

Table 2 | Na<sup>+</sup>-binding site integrity in different states of BetP measured as the distance between coordinating groups in X-ray crystallographic structures and during molecular dynamics simulations thereof.

State	System	d(Na1')	$\Delta d(Na1')$	d(Na2)	$\Delta d(Na2)$
C <sub>c</sub> S	4AIN-B	4.7	NA	3.8	NA
	MD*	$4.7 \pm 0.3$	0.0	$3.5 \pm 0.2$	-0.3
Cc	4AIN-A	4.8	0.1	5.0	1.2
C <sub>e</sub> S	4LLH-A	4.9	0.2	4.7	0.9
	4LLH-B	4.7	0.0	3.7	-0.1
	MD	$5.8 \pm 0.5$	1.1	$4.0 \pm 0.5$	0.2
C <sub>e</sub>	4DOJ-B	6.7	2.0	5.9	2.1
	MD	$5.8 \pm 0.4$	1.1	$6.1 \pm 1.4$	2.3
C <sub>i</sub> S	4DOJ-C	8.0	3.3	7.7	3.9
	4LHH-C	8.0	3.3	7.5	3.7
	3PO3-C	8.3	3.6	7.4	3.6
$C_{i}$	3PO3-A	8.9	3.6	6.9	3.1
	3PO3-B	6.8	2.1	7.3	3.5

NA, not applicable.

Systems are either crystal structures, listed as PDB and chain identifier, or mean ± s.d. computed over three 100-ns molecular dynamics simulations of those structures, where available.
\*Simulations of C<sub>c</sub>S state described previously<sup>10</sup>. Ion-binding sites are defined as follows: for Na1, the distance (in Å) between T250 side-chain oxygen (TMH3') and the center of the F380 (TMH5') phenyl ring; for Na2, the distance between the carbonyl oxygen atoms of M150 (TMH1') and F464 (TMH8'). The difference in binding site distances (\(\Delta d\)) is calculated relative to the value for the C<sub>c</sub>S state in 4.HN.B.

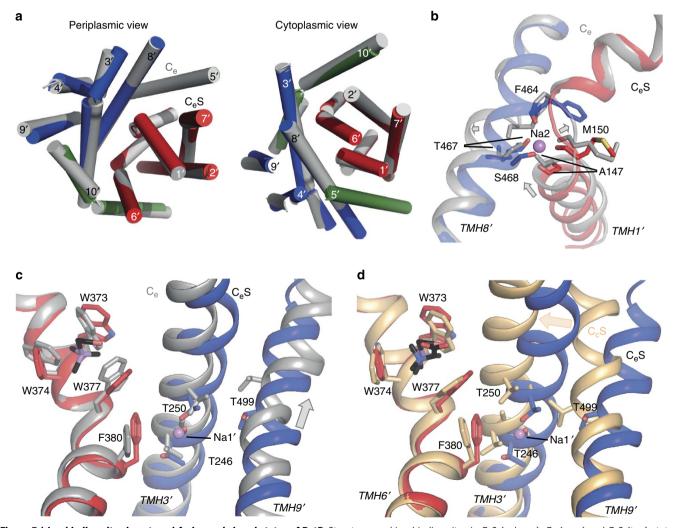


Figure 5 | Ion-binding sites in outward-facing and closed states of BetP. Structures and ion-binding sites in  $C_eS$  (colours),  $C_e$  (grey) and  $C_cS$  (tan) states of BetP and BetP-G153D are compared. Sodium ions (violet spheres), and substrate (black sticks) are shown. Changes from the  $C_eS$  state to the compared state are indicated using arrows. ( $\mathbf{a}$ - $\mathbf{c}$ ) Comparisons of substrate-free  $C_e$  (PDB entry 4DOJ chain B, grey) and substrate-bound  $C_eS$  states (colours), showing the effect of substrate binding on the outward-open states on: ( $\mathbf{a}$ ) the overall structure, with helices shown as cylinders; ( $\mathbf{b}$ ) the Na2-binding site, showing protein (ribbons), binding site residues (sticks) and Na2 ion from the  $C_eS$  structure (sphere). ( $\mathbf{c}$ ) The substrate site (S1) and Na1' site, with protein shown as ribbons, and with substrate (black) and binding-site residues as sticks. The location of the Na1' ion is taken from the  $C_eS$  simulations.

( $\mathbf{d}$ ) Comparison of Na1' and substrate (S1)-binding site regions in the substrate-bound outward-open ( $C_eS$ , colours) and the substrate-bound closed ( $C_cS$ ; PDB entry 4AIN chain B, tan) states. The Na1' ion location is taken from the  $C_eS$  simulations.

includes two water molecules in the  $C_eS$  state (Fig. 6b) or a single water in the  $C_rS$  state  $^{10}$ .

In spite of the tenuous and minimal number of direct interactions with the protein, the sodium ion at Na1' in the  $C_e$  state formed a stable network of interactions with  $3.4\pm0.5$  water molecules (Figs 6d,e and 7). These water molecules in turn formed a network of hydrogen bonds with surrounding groups including, notably, the side chain amine of Trp377 from TMH6' and the Ser253 hydroxyl from TMH3', both of which shape the substrate-binding site². The hydrogen bond network also occasionally included the backbone carbonyl of Ala469 or the hydroxyl of Ser468 from TMH8', which forms the Na2 site (Fig. 6d).

We reiterate that no positive peak in the Fo-Fc electron-density map at the Na1' site was observed in the C<sub>e</sub>S state nor in any of the previously crystallized states of BetP<sup>2</sup>, despite strong evidence that an ion binds to that region, as described above. This lack of a

peak might plausibly be a consequence of the dynamic nature of such a well-hydrated ion-binding site, or potentially also of a reduction in occupancy at the Na1' site in crystals upon dehydration.

#### DISCUSSION

To date, structures of BetP WT and the choline-transporting mutant BetP-G153D have provided eight distinct conformational states (C<sub>eoc</sub>, C<sub>e</sub>Na, C<sub>e</sub>S, C<sub>c</sub>S, C<sub>ioc</sub>S, C<sub>i</sub>S, C<sub>i</sub> and C<sub>c</sub>) (Fig. 8, Supplementary Movies 1–3)<sup>2,8,23</sup>, which after comparison with the crystallized outward- and inward-facing states of Mhp1 and LeuT, revealed similar mechanistic principles but also unique features in each individual system<sup>2</sup>. The choline-transporting BetP-G153D exhibits a higher affinity for Na <sup>+</sup> than WT protein. This mutant also tends to crystallize in the apo outward-occluded and outward-open states, while the WT crystallizes in either

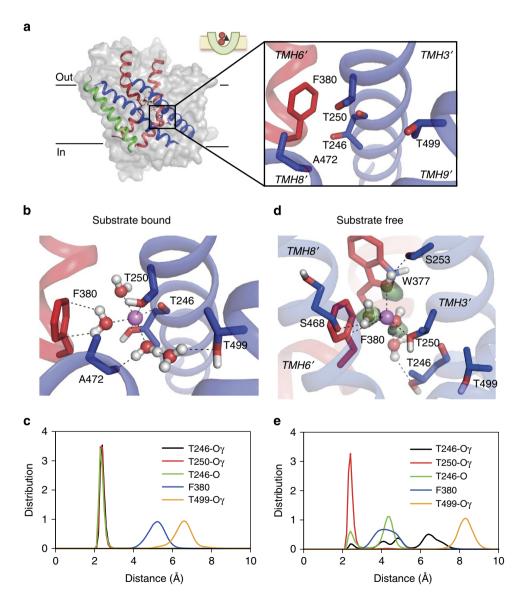


Figure 6 | The Na1'-binding site in outward-open states of BetP-G153D. (a) The  $C_eS$  state structure (protomer B) in side view with Na1' site and adjacent central substrate-binding site in stick representation, coordinating arseno-choline (black, red and purple sticks). The inset shows residues that form the Na1' site as reported for the closed substrate-bound  $C_cS$  state. (b-e) The Na1' site in simulations of the substrate-bound  $C_eS$  (b,c) and substrate-free  $C_e$  (d-e) states: (b,d) example configurations in simulations of the substrate-bound  $C_eS$  state (d). Proteins and ions are presented as in Fig. 6. (c,e) Distances between the Na<sup>+</sup> ion and protein-binding site atoms (either oxygen, or the center of the phenyl ring in F380) during simulations of the substrate-bound  $C_eS$  state (c) and the substrate-free  $C_e$  state (e).

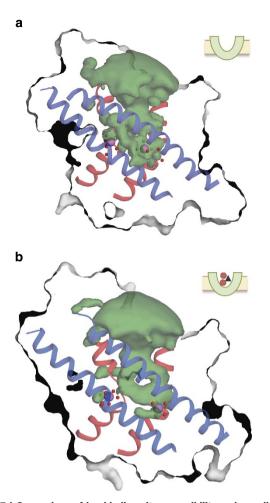


Figure 7 | Comparison of ion-binding site accessibility and coordination in the substrate-free  $C_e$  (a) and substrate-bound  $C_eS$  (b) states of BetP-G153D observed using molecular dynamics simulations. The green surface indicates the density of water molecules occupying the extracellular pathway (defined as waters within 16 Å of D153) during three 200 ns trajectories. Protein snapshots were taken at 81 and 83 ns of the trajectories of the  $C_e$  and  $C_eS$  states, respectively. BetP protein is shown as a cut-away surface. Sodium ions are shown as violet spheres. Oxygen atoms of water and protein that form the Na2- and Na1'-binding site in the closed  $C_cS$  state are shown as red spheres. Oxygen atoms in addition to those from the closed-state structure coordinating the ion during molecular simulations of the corresponding state are shown as yellow spheres. Selected helices are shown as cartoons: TMH1' and its symmetry equivalent TMH6' (red), and TMH3' and its symmetry equivalent TMH8' (blue), with substrate choline shown as sticks.

inward or closed states, depending on the concentration of the cocrystallized substrate. Thus, we have now been able to observe substrate in all three main chain conformations, namely  $C_eS$ ,  $C_cS$ and  $C_iS$ . This wealth of BetP structures reveals distinct positions of the substrate during its transport trajectory through BetP. While it should be kept in mind that we are comparing structures that have been crystallized with different substrates, we note that the inward-facing  $C_iS$  state has been observed with both betaine and choline, with nearly identical positions for both trimethyammonium compounds<sup>2</sup>, and therefore the distinct positions observed for substrate in the different states are likely to be physiologically relevant.

The anomalous scattering demonstrated that arseno-choline is bound to the structure reported here, which allowed a

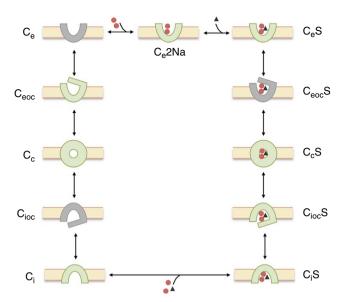


Figure 8 | The eight conformations of BetP and BetP-G153D determined by X-ray crystallography to date (green): substrate-free outward-open bound to sodium ( $C_e$ 2Na) (PDB entry 4DOJ); substrate-bound outward-open ( $C_e$ S) (PDB entry 4LLH, reported here); substrate-bound closed ( $C_c$ S) (PDB entry 4AIN); substrate-bound inward-occluded ( $C_{ioc}$ S) (PDB entry 2WIT); substrate-bound inward-open ( $C_i$ S) (PDB entries 4DOJ/3PO3/4AIN/4LLH); substrate-free inward-open ( $C_i$ ) (PDB entry 4AMR); substrate-free closed ( $C_c$ ) (PDB entry 4AIN); and substrate-free outward-occluded ( $C_{eoc}$ ) (PDB entry 4DOJ).

comparison of substrate-free (Ce) and -bound (CeS) outwardopen states. MD simulations suggest that sodium is bound, albeit loosely, to both the Na1' and Na2 sites in both states (Figs 4 and 6), and therefore we reassign the Ce structure (monomer B of Protein Data Bank (PDB) entry 4DOJ) to a C<sub>e</sub>2Na state (Fig. 8). The observation that an ion remains coordinated at a solventaccessible Na2 site in the substrate-free Ce2Na state mirrors observations for an outward-open structure of another LeuT-fold transporter, Mhp1 (ref. 6). In both transporters, binding of substrate and formation of the CeS state leads to occlusion of the Na2 site from the extracellular solution (Fig. 7)<sup>6</sup>. Our simulations also suggest a concomitant increase in stability of the ion-protein interactions at the Na2 site in the CeS state compared with the C<sub>e</sub>2Na state (Fig. 4). These changes reflect a number of small helix reorganizations around the site that improve the Na2 coordination in the  $C_eS$  state (Fig. 5b), rendering it very similar to that in the  $C_eS$  state (Table 2)<sup>10</sup>.

In the case of the Na1′ site, by contrast, the ion is accessible to water in both the  $C_e2Na$  and  $C_eS$  states (Fig. 7). However, similar to Na2, the Na1′ site exhibits significantly fewer stable protein–ion interactions in the  $C_e2Na$  state compared with the substrate-bound form ( $C_eS$ , Fig. 6) or the closed  $C_cS$  state  $^{10}$ , due to the shift in TMH9′ (Fig. 5C). Complete closure of the periplasmic pathway in the transition between the two holo forms ( $C_eS$  to  $C_cS$ ) involves further contraction of TMH3′ and TMH9′ towards TMH8′ (Fig. 5d), which brings the ion  $\sim 0.5$  Å closer to F380 (Fig. 6 cf. (ref. 10)).

Meta-stable ion-binding sites, such as those seen here for the  $C_e2Na$  state, that are proximal to, and involve a number of ligands from, the main binding sites, may be useful because they enhance the local ion concentration. Indeed, solvated, meta-stable sites appear to be a common theme for ion-coupled transporters, and have been identified during MD simulations and free-energy calculations of  $vSGLT^{20,21}$  and  $Mhp1^{24}$ , as well as  $LeuT^{25}$ .

The important role of individual water molecules binding sodium within transporter proteins is underlined by the ion-water coordination observed in a recent structure of *Drosophila melanogaster* dopamine transporter (dDAT)<sup>26</sup>, as well as in c-subunits of ATP-synthase rotor rings<sup>27</sup>.

Together, these results suggest a sequential Na+/substrate-binding process in which ions are loosely associated with partially formed, partially hydrated Na1' and Na2 sites in the outward-open state, in the absence of substrate. Studies of other LeuT homologues indicate distinct roles for the two cations in transport<sup>28–32</sup> and, in terms of binding order, the most likely scenario appears to be that Na2 would bind before Na1<sup>28</sup>. Indeed, spectroscopic and computational evidence that sodium shifts the equilibrium of LeuT towards opening of the central binding site to the periplasm<sup>11,12,24</sup>; this, combined with the fact that the Na2 site is conserved among sodium-coupled transporters, suggests that Na2 is required for the conformational change. A similar mechanism may be adopted by BetP since the Na2 site residues are closer together in the outward-open states than in the inward-open states (Table 2).

The finding that both ions bind to the empty transporter before substrate is consistent with electrophysiological studies of human Na<sup>+</sup>-coupled glucose symporter (hSGLT1), a homologue of vSGLT<sup>30</sup>. In the case of BetP and Mhp1, one effect of substrate binding is to occlude the Na2 site, which should reduce the probability of ion unbinding<sup>6,24</sup>. Complete closure to form the C<sub>c</sub>S state then involves optimization of the binding interactions, at which point the energy gained by binding is maximal. By analogy with enzymatic mechanisms, this closed state is therefore a transition state, in line with the induced transition fit model proposed by Klingenberg<sup>33</sup>. However, further analysis of BetP will be required to further validate this hypothesis. For example, in future work it would be of interest to follow the approach of Loo *et al.*<sup>30</sup>, and to compute the potential of mean force of the ions along their pathway into the binding sites in BetP to predict whether there is a significant barrier to formation of these assumed states.

In conclusion, results from a combination of structural studies and molecular simulations reveal for the first time for a secondary transporter how binding of coupling ions may involve progressive dehydration and concurrent coordination by protein residues. Studies of subsequent steps in the cycle will be required to address the question of whether the binding steps alone, or other later events, determine the strict ion-substrate coupling in BetP.

# Methods

**Synthesis of arseno-choline.** Arseno-choline was obtained<sup>34</sup> by heating trimethylarsine with bromo-ethanol, followed by re-crystallization from methanol/acetone; the purity (>99.5%) of the compounds was established by NMR spectroscopy and HPLC/mass spectrometry (MS) (inductively coupled plasma MS and electrospray MS).

**Cell culturing and protein purification.** Cell culture and protein preparation methods have been described previously<sup>23</sup>. Briefly, *E. coli* DH5αmcr<sup>35</sup> were use for the heterologous expression of strep-betP. Cells were grown at 37 °C in LB medium supplemented with carbenicillin (50 μg ml - 1) and induction was initiated with anhydrotetracycline (200 µg l<sup>-1</sup>). Cells were harvested at 4 °C by centrifugation and resuspended in buffer containing 100 mM Tris-HCl (pH 8.0) and protease inhibitor Pefabloc 0.24 mg ml - 1. Membranes were isolated from disrupted cells and solubilized with 1%  $\beta$ -dodecyl-maltoside (DDM) when purified protein was subsequently crystallized or reconstituted. The protein was then loaded on a StrepII-Tactin macroprep column, washed with 50 mM Tris-HCl (pH 7.5), 500 mM NaCl, 8.6% glycerol, 0.1% DDM and eluted with 5 mM desthiobiotin, 50 mM Tris-HCl (pH 7.5), 200 mM NaCl, 8.7% glycerol and 0.6% Cymal-5, if used for crystallization, or 0.1% DDM if used for reconstitution. Before crystallization the protein was loaded onto a Superose 6 (GE Healthcare) size-exclusion column equilibrated with 20 mM Tris-HCl (pH 7.5), 200 mM NaCl and 0.6% Cymal-5. This purified protein was concentrated at  $4\,^{\circ}\text{C}$  to  $\sim 10\,\text{mg}\,\text{ml}^{-1}$  at  $3{,}000\,\text{g}$  in a Vivaspin tube (Vivascience) with a 100-k-molecular-weight cutoff, and incubated for 16 h at 4 °C with either 8 mM arseno-choline or 1 mM arseno-betaine.

Site-directed mutagenesis. Site-directed mutagenesis of the pASK-IBA5betP20 and pASK IBA7betP $\Delta$ N29/E44E45E46/AAA23 plasmids were performed using the QuickChangeTM kit (Stratagene), Pfu Turbo DNA polymerase and specific Oligonucleotides. The mutations were verified by nucleotide sequencing. The fully functional mutant BetP( $\Delta$ N29/E44E45E46/AAA/G153D) was used for crystallization purposes yielding in improved diffraction power and isotropy of the crystals.

Crystallization and structure determination. BetP- $\Delta$ N29/E44E45E46/AAA/G153D was co-crystallized with 8 mM arseno-choline at 18 °C by vapour diffusion in hanging drops of 1 µl protein solution with 1 µl of 100 mM Na-tri-citrate (pH 5.3–5.6), 300 mM NaCl or RbCl and 17–24% PEG 400 as reservoir solution. Crystals of BetP- $\Delta$ N29/EEE44/45/46AAA/G153D diffracted to 2.9 Å, and data were collected on the beamline PXII at the Swiss Light Source (Swiss Light Source at the Paul Scherrer Institut, Villigen, Switzerland). Data were collected at the K edge for arsenic, using a wavelength of  $\sim$ 1.043 Å. Data were processed using the XDS package<sup>36</sup> and the anisotropy was corrected on the UCLA Diffraction Anisotropy Server (http://services.mbi.ucla.edu/anisoscale/)<sup>37</sup>. Structure was determined by molecular replacement with BetP (PDB entry 4DOI) as search model using PHASER<sup>38</sup> and refinement was performed using PHENIX<sup>39</sup> combined with manual rebuilding in COOT<sup>40</sup>.

**Protein reconstitution into liposomes.** Functional reconstitution of BetP-G153 was performed as described<sup>41</sup>. Briefly, liposomes (20 mg phospholipid ml $^{-1}$ ) from *E. coli* polar lipids (Avanti polar lipids) were prepared by extrusion through polycarbonate filters (100 nm pore size) and diluted 1:4 in buffer (250 mM KPi pH 7.5). After saturation with Triton X-100, the liposomes were mixed with purified protein at a lipid/protein ratio of 10:1 (w/w). BioBeads at ratios (w/w) of 5 (BioBeads/Triton X-100) and 10 (BioBeads/DDM) were added to remove detergent. Finally, the proteoliposomes were centrifuged and washed before being frozen in liquid nitrogen and stored at - 80  $^{\circ}$ C.

**Tryptophan fluorescence-binding assay.** Binding assays were performed with  $100\,\mu g\,ml^{-1}$  of purified BetP-G153D in proteoliposomes  $^{41}$ . Choline and arsenocholine concentrations ranged from 0.1 to  $16\,mM$  for measurements at  $100\,mM$  NaCl and pH 7.5. Choline concentrations ranged from 0.8 to  $30\,mM$  at pH 5.5 and 0.8 to  $100\,mM$  at pH 7.5 in absence of NaCl. Tryptophan fluorescence emission between 315 and 370 nm was recorded on a Hitachi F-4500 fluorescence spectrophotometer and averaged over eight readings, with the excitation wavelength set to  $295\,nm$  and a slit width of 2.5 or 5.0 nm for excitation or emission, respectively. The mean value and s.d. at the 340-nm emission maximum was plotted for each substrate concentration. Binding constants were derived by fitting with the program GraphPad Prism version 5.0c for Mac OS X, GraphPad Software  $^{42}$ .

**SSM-based electrophysiology**. SSM-based electrophysiology was performed as described  $^{10,43,44}$  using  $40\,\mu l$  of proteoliposomes at a protein concentration of 1 mg ml <sup>-1</sup>. Briefly, experiments were carried out at room temperature (22 °C). The solution-exchange protocol consisted of four phases with a total duration 4.8 s. (i) non-activating solution (2.5 s), (ii) activating solution (0.8 s), (iii) non-activating solution (0.5 s) and (iv) resting solution (1.0 s). The resting solution was composed of 250 mM KPi buffer (7.5). Non-activating and activating solution also contained 250 mM KPi buffer (pH 7.5). The osmolality of non-activating and activating solution was kept constant. Activation by choline or arseno-choline was performed using a non-activating solution with x µM glycine plus 500 mM NaCl and an activating solution with corresponding x µM choline or arseno-choline plus 500 mM NaCl. Thereby an osmolar gradient was established when switching from the resting to the non-activating solution at the start of the experiment. Note that during the final experiments (several minutes) the cuvette contained resting solution to ensure low osmolar conditions in the proteoliposomes. Transient currents were recorded at the concentration jumps taking place from nonactivating to activating buffers. The current amplifier was set to a gain of  $10^9$ –  $10^{10}\,\mathrm{V\,A^{-1}}$ , and low-pass filtering was a 300–1,000 Hz.

**Transport assays.** Uptake of [ $^{14}$ C]-betaine in *E. coli* cells was performed as described  $^{45}$ . *E. coli* MKH13 cells expressing a particular strep-betP mutant were cultivated at 37  $^{\circ}$ C in LB medium containing carbenicillin (50 µg ml  $^{-1}$ ) and induced at an  $OO_{600}$  of 0.5 by adding anhydrotetracycline (200 µg l  $^{-1}$ ). After 2 ht he cells were harvested and washed in buffer containing 25 mM KPi buffer (pH 7.5) and then were resuspended in the same buffer containing 20 mM glucose. For uptake measurements the external osmolality was adjusted with KCl at a constant value of 800 mOsmol kg  $^{-1}$ . Sodium titration was performed by adjusting the NaCl concentration in the buffer. Cells were incubated for 3 min at 37  $^{\circ}$ C before the addition of 250 µM [ $^{14}$ C]-betaine. Betaine uptake was measured at various time intervals, after cell samples were passed through glass fibre filters (APFF02500; Millipore) and were washed twice with 2.5 ml of 0.6 M KPi buffer. The radioactivity retained on the filters was quantified by liquid scintillation counting.

**MD simulations.** MD simulations were carried for monomers taken from two different crystallographic structures: (1) monomers A and B of the structure presented here (PDB 4LLH), which are both in outward-open substrate-bound ( $C_e$ S) states; and (2) monomer B of PDB entry 4DOJ, which is in an outward-open apo ( $C_e$ ) state. Protonation states were assigned according to calculations performed on the X-ray structures using MCCE2.0<sup>46</sup>. Specifically, residues His188, His192, His445, His482 and His570 were protonated at their Nɛ atom while His176, His176 and His573 were protonated at their Nɛ atoms. On the basis of the same calculations, Glu161 was assumed to be neutral, while Asp153 was charged. The likely location of buried water molecules was identified using DOWSER<sup>47</sup>.

The protein was inserted into a 50:50 diastereoisomeric mixture of 219 palmitoyl oleoyl phosphatidyl-glycerol (POPG) lipid molecules, hydrated with salt water. Specifically, monomers A and B of the  $C_eS$  state were solvated by 15,140 and 15,149 water molecules, plus 245 and 254 sodium ions, respectively, and the  $C_e$  structure was solvated by 15,126 water molecules and 253 sodium ions. All there e systems also contain 27 chloride ions to neutralize any net charge and to maintain an  $\sim$  0.15 M salt concentration. The final systems were  $\sim$  91  $\times$  91  $\times$  96 Å in size.

In the  $C_eS$  structures, a substrate choline molecule and an ion at Na2 were placed according to positive peaks observed in the  $F_o-F_c$  difference-density map, whereas the sodium at Na1′ was placed equidistant between the hydroxyl oxygen atoms of T246 and T250 and the center of the F380 phenyl ring. A water molecule was placed adjacent to the Na1′ so as to indirectly coordinate T499 (ref. 10). In the  $C_e$  structure, sodium ions were placed based on superposition of the simulated  $C_cS$  structure (monomer B of 4AIN).

The complete monomer-bilayer system was set up using GRIFFIN<sup>48</sup>. Briefly, GRIFFIN was first used to carve out lipid and water molecules from the equivalent volume of the protein. Subsequently, a three-step optimization was performed, with each step 50 ps long and with external forces of 1.0, 2.0, then 3.0 kcal mol  $^{-1}$  Ų or 1.0, 1.5 then 2.0 kcal mol  $^{-1}$  Ų used for the  $\rm C_e$  or  $\rm C_e$ S structures, respectively. After 1,000 steps of conjugate gradients energy minimization, 20 ns of MD simulation was carried out in which all non-hydrogen atoms of the protein, ligand and bound sodium ions were constrained to their initial positions using springs with progressively smaller force constants, starting at 15 kcal mol  $^{-1}$  Ų. During the equilibration of the  $\rm C_e S$  state, distance constraints were applied between choline and the tryptophan box (W373, W374 and W377) as well as with D153, in order to maintain the substrate orientation and position as in the  $\rm C_e S$  X-ray structures. Similarly, the distances between sodium ions and their coordinating residues (T246, T250 and F380 for Na1 and A147, M150, F464, T467 and S468 for Na2) were constrained. Analysis was carried out every ps of unconstrained simulations, each 200 ns (for  $\rm C_e S$ ) or 100 ns (for  $\rm C_e$ ) in length.

All MD simulations were performed using NAMD<sup>49</sup>. Periodic boundary conditions were used. A real-space cutoff of 12 Å was used for both van der Waals and long-range electrostatics; the distance at which the switching function began to take effect was 10 Å. The time step was 2 fs. The SHAKE algorithm<sup>50</sup> was used to fix all bond lengths. Constant temperature (310 K) was set with a Langevin thermostat<sup>51</sup>, with a coupling coefficient of 0.2 ps $^{-1}$ . A Nosé–Hoover Langevin barostat $^{52}$  was used to apply constant pressure normal to the membrane plane, with an oscillation period of 200 fs and the damping time scale set to 50 fs. The surface area in the membrane plane was kept constant.

The all-atom CHARMM27 force field was used for protein<sup>53,54</sup> and ions<sup>55</sup>, and TIP3P was used for water molecules<sup>56</sup>. Force field parameters for POPG were kindly provided by H. Jang (National Cancer Institute, Frederick, MD)<sup>57</sup>. The parameters for choline were taken from the CHARMM general force field (CGenFF) topology and parameter files<sup>58</sup>.

The MD trajectories were analysed with CHARMM and visual MD<sup>59</sup>.

# References

- Yamashita, A., Singh, S. K., Kawate, T., Jin, Y. & Gouaux, E. Crystal structure of a bacterial homologue of Na<sup>+</sup>/Cl<sup>-</sup>-dependent neurotransmitter transporters. *Nature* 437, 215–223 (2005).
- Perez, C., Koshy, C., Yildiz, O. & Ziegler, C. Alternating-access mechanism in conformationally asymmetric trimers of the betaine transporter BetP. *Nature* 490, 126–130 (2012).
- 3. Forrest, L. R., Kramer, R. & Ziegler, C. The structural basis of secondary active transport mechanisms. *Biochim. Biophys. Acta* **1807**, 167–188 (2011).
- Forrest, L. R. et al. Mechanism for alternating access in neurotransmitter transporters. Proc. Natl Acad. Sci. USA 105, 10338–10343 (2008).
- Wong, F. H. et al. The amino acid-polyamine-organocation superfamily. J. Mol. Microbiol. Biotechnol. 22, 105–113 (2012).
- Micropiol. Biolectinol. 22, 105–113 (2012).
   Shimamura, T. et al. Molecular basis of alternating access membrane transport by the sodium-hydantoin transporter Mhp1. Science 328, 470–473 (2010).
- Faham, S. et al. The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na+/sugar symport. Science 321, 810–814 (2008).
- 8. Perez, C. et al. Substrate specificity and ion coupling in the Na<sup>+</sup>/betaine symporter BetP. EMBO J. **30**, 1221–1229 (2011).
- Francesconi, K. A. Current perspectives in arsenic environmental and biological research. *Environ. Chem.* 2, 141–145 (2005).
- Khafizov, K. et al. Investigation of the sodium-binding sites in the sodium-coupled betaine transporter BetP. Proc. Natl Acad. Sci. USA 109, E3035–E3044 (2012).

- Zhao, Y. et al. Single-molecule dynamics of gating in a neurotransmitter transporter homologue. Nature 465, 188–193 (2010).
- Claxton, D. P. et al. Ion/substrate-dependent conformational dynamics of a bacterial homolog of neurotransmitter:sodium symporters. Nat. Struct. Mol. Biol. 17, 822–829 (2010).
- Kowalczyk, L. et al. Molecular basis of substrate-induced permeation by an amino acid antiporter. Proc. Natl Acad. Sci. USA 108, 3935–3940 (2011).
- Singh, S. K., Yamashita, A. & Gouaux, E. Antidepressant binding site in a bacterial homologue of neurotransmitter transporters. *Nature* 448, 952–956 (2007).
- Oswald, C. et al. Crystal structures of the choline/acetylcholine substratebinding protein ChoX from Sinorhizobium meliloti in the liganded and unliganded-closed states. J. Biol. Chem. 283, 32848–32859 (2008).
- Fernandez-Tornero, C., Lopez, R., Garcia, E., Gimenez-Gallego, G. & Romero, A. A novel solenoid fold in the cell wall anchoring domain of the pneumococcal virulence factor LytA. Nat. Struct. Biol. 8, 1020–1024 (2001).
- Pittelkow, M., Tschapek, B., Smits, S. H., Schmitt, L. & Bremer, E. The crystal structure of the substrate-binding protein OpuBC from Bacillus subtilis in complex with choline. J. Mol. Biol. 411, 53–67 (2011).
- Perez, C. & Ziegler, C. Mechanistic aspects of sodium-binding sites in LeuT-like fold symporters. *Biol. Chem.* 394, 641–648 (2013).
- Kalayil, S., Schulze, S. & Kühlbrandt, W. Arginine oscillation explains Na independence in the substrate/product antiporter CaiT. *Proc. Natl Acad. Sci. USA* 110, 17296–17301 (2013).
- Bisha, I., Laio, A., Magistrato, A., Giorgetti, A. & Sgrignani, J. A candidate ionretaining state in the inward-facing conformation of sodium/galactose symporter: clues from atomistic simulations. *J. Chem. Theory. Comput.* 9, 1240–1246 (2013).
- 21. Li, J. & Tajkhorshid, E. Ion-releasing state of a secondary membrane transporter. *Biophys. J.* **97**, L29–L31 (2009).
- Watanabe, A. et al. The mechanism of sodium and substrate release from the binding pocket of vSGLT. Nature 468, 988–991 (2010).
- Ressl, S., Terwisscha van Scheltinga, A. C., Vonrhein, C., Ott, V. & Ziegler, C. Molecular basis of transport and regulation in the Na(+)/betaine symporter BetP. *Nature* 458, 47–52 (2009).
- 24. Zhao, C. & Noskov, S. Y. The molecular mechanism of ion-dependent gating in secondary transporters. *PLoS Comput. Biol.* **9**, e1003296 (2013).
- Zhao, C. et al. Ion-controlled conformational dynamics in the outward-open transition from an occluded state of LeuT. Biophys. J. 103, 878–888 (2012).
- Penmatsa, A., Wang, K. H. & Gouaux, E. X-ray structure of dopamine transporter elucidates antidepressant mechanism. *Nature* 503, 85–90 (2013).
- Meier, T. et al. Complete ion-coordination structure in the rotor ring of Na<sup>+</sup>-dependent F-ATP synthases. J. Mol. Biol. 391, 498–507 (2009).
- Meinild, A.-K. & Forster, I. C. Using lithium to probe sequential cation interactions with GAT1. Am. J. Physiol. Cell Physiol. 302, C1661–C1675 (2012).
- Felts, B. et al. The two Na<sup>+</sup> sites in the human serotonin transporter play distinct roles in the ion coupling and electrogenicity of transport. J. Biol. Chem. 289, 1825–1840 (2013).
- Loo, D. D. F., Jiang, X., Gorraitz, E., Hirayama, B. A. & Wright, E. M. Functional identification and characterization of sodium binding sites in Na symporters. *Proc. Natl Acad. Sci. USA* 110, E4557–E4566 (2013).
- Noskov, S. Y. & Roux, B. Control of ion selectivity in LeuT: two Na<sup>+</sup> binding sites with two different mechanisms. J. Mol. Biol. 377, 804–818 (2008).
- Zhou, Y., Zomot, E. & Kanner, B. I. Identification of a lithium interaction site in the gamma-aminobutyric acid (GABA) transporter GAT-1. *J. Biol. Chem.* 281, 22092–22099 (2006).
- Klingenberg, M. Ligand-protein interaction in biomembrane carriers. The induced transition fit of transport catalysis. *Biochemistry* 44, 8563–8570 (2005).
- Irgolic, K. J., Junk, T., Kos, C., McShane, W. S. & Pappalardo, G. C. Preparation of trimethyl-2-hydroxyethylarsonium (arsenocholine) compounds. *Appl. Organometal. Chem.* 1, 403–412 (1987).
- Grant, S. G., Jessee, J., Bloom, F. R. & Hanahan, D. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc. Natl Acad. Sci. USA* 87, 4645–4649 (1990).
- Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J. Appl. Cryst. 26, 795–800 (1993).
- Strong, M. et al. Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from Mycobacterium tuberculosis. Proc. Natl Acad. Sci. USA 103, 8060–8065 (2006).
- McCoy, A. J. et al. Phaser crystallographic software. J. Appl. Cryst. 40, 658–674 (2007).
- Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta. Cryst. D 66, 213–221 (2010).
- Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta. Cryst.allogr. D Biol. Crystallogr.* 60, 2126–2132 (2004).
- 41. Ge, L., Perez, C., Waclawska, I., Ziegler, C. & Muller, D. J. Locating an extracellular K + -dependent interaction site that modulates betaine-binding of

- the Na <sup>+</sup>-coupled betaine symporter BetP. *Proc. Natl Acad. Sci. USA* **108**, E890–E898 (2011).
- Motulsky, H. Analyzing Data with GraphPad Prism (GraphPad Software Inc., 1999) http://www.graphpad.com.
- Schulz, P., Garcia-Celma, J. J. & Fendler, K. SSM-based electrophysiology. *Methods* 46, 97–103 (2008).
- Khafizov, K. et al. Structural symmetry reveals the sodium binding sites in the betaine transporter BetP. Proc. Natl Acad. Sci. USA 109, E3035–E3044 (2011).
- Perez, C., Khafizov, K., Forrest, L. R., Kramer, R. & Ziegler, C. The role of trimerization in the osmoregulated betaine transporter BetP. EMBO. Rep. 12, 804–810 (2011).
- Alexov, E. G. & Gunner, M. R. Incorporating protein conformational flexibility into the calculation of pH-dependent protein properties. *Biophys. J.* 72, 2075–2093 (1997).
- Zhang, L. & Hermans, J. Hydrophilicity of cavities in proteins. *Proteins* 24, 433–438 (1996).
- Staritzbichler, R., Anselmi, C., Forrest, L. R. & Faraldo-Gomez, J. D. GRIFFIN: a versatile methodology for optimization of protein-lipid interfaces for membrane protein simulations. J. Chem. Theory Comput. 7, 1167–1176 (2011).
- Phillips, J. C. et al. Scalable molecular dynamics with NAMD. J. Comput. Chem. 26, 1781–1802 (2005).
- Ryckaert, J.- P., Ciccotti, G. & Berendsen, H. J. C. Numerical-integration of cartesian equations of motion of a system with constraints—molecular dynamics of n-alkanes. *J. Comput. Phys.* 23, 327–341 (1977).
- Adelman, S. A. & Doll, J. D. Generalized Langevin equation approach for atomsolid-surface scattering: general formulation for classical scattering off harmonic solids. J. Chem. Phys. 64, 2375–2388 (1976).
- Feller, S. E., Zhang, Y. H., Pastor, R. W. & Brooks, B. R. Constant-pressure molecular-dynamics simulation: the Langevin piston method. *J. Chem. Phys.* 103, 4613–4621 (1995).
- 53. Mackerell, Jr A. D., Feig, M. & Brooks, 3rd C. L. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. J. Comput. Chem. 25, 1400–1415 (2004).
- Feller, S. E., Gawrisch, K. & MacKerell, Jr A. D. Polyunsaturated fatty acids in lipid bilayers: intrinsic and environmental contributions to their unique physical properties. J. Am. Chem. Soc. 124, 318–326 (2002).
- Nina, M., Beglov, D. & Roux, B. Atomic radii for continuum electrostatics calculations based on molecular dynamics free energy simulations. *J. Phys. Chem. B* 101, 5239–5248 (1997).
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79, 926–935 (1983).

- Jang, H., Ma, B., Woolf, T. B. & Nussinov, R. Interaction of protegrin-1 with lipid bilayers: membrane thinning effect. *Biophys. J.* 91, 2848–2859 (2006).
- Vanommeslaeghe, K. et al. CHARMM general force field: a force field for druglike molecules compatible with the CHARMM all-atom additive biological force fields. J. Comput. Chem. 31, 671–690 (2010).
- Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. J. Mol. Graph. 14, 33–38 (1996).

## Acknowledgements

We thank Susanne Ressl for important insights on using arseno-compounds for cocrystallization, and Caroline Koshy and José Faraldo-Gómez for helpful discussions. We thank Klaus Fendler and Özkan Yildiz for their support with SSM and X-ray diffraction experiments. We also thank the staff of the Swiss Light Source (SLS, PXII) for their continuous support. This work was supported by the International Max-Planck Research School (IMPRS), the Deutsche Forschungsgemeinschaft (SFB807 'Transport and communication across biological membranes') and the Transport Across Membranes grant (TRAM).

# **Author contributions**

C.P. performed transport measurements, processing and refinements of crystallographic data; B.F. performed transport and binding measurements, collection and processing of data; A.R.M. performed molecular dynamics simulations; K.A.F. performed synthesis of arseno-compounds; B.F., A.R.M., L.R.F., C.Z. and C.P. analysed the data; C.P., L.R.F. and C.Z. directed the research; and C.P., L.R.F. and C.Z. wrote the manuscript.

#### **Additional information**

Accession codes. Coordinates and structure factors for the outward-open state of BetP have been deposited in the RCSB Protein Data Bank under accession code 4LLH.

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://www.nature.com/ reprints/index.html.

**How to cite this article:** Perez, C. *et al.* Substrate-bound outward-open state of the betaine transporter BetP provides insights into Na <sup>+</sup> coupling. *Nat. Commun.* 5:4231 doi: 10.1038/ncomms5231 (2014).